VALUE IN HEALTH 14 (2011) 1010–1018

Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Cost-Utility Analyses of Diagnostic Laboratory Tests: A Systematic Review

ChiHui Fang, MPH, MS[1], Hansel J Otero, MD[2], Dan Greenberg, PhD[1,3], Peter J Neumann, ScD[1,]*

[1]*Center for the Evaluation of Value and Risk in Health, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA, USA;* [2]*Department of Radiology, Tufts Medical Center, Boston, MA, USA;* [3]*Department of Health Systems Management, Ben-Gurion University of the Negev, Beer-Sheva, Israel*

ABSTRACT

Objective: To review and evaluate the literature of cost-utility analyses (CUAs) regarding diagnostic laboratory testing. **Methods:** We reviewed all articles related to diagnostic laboratory testing in the Tufts Medical Center Cost-Effectiveness Analysis Registry (www.cearegistry.org), which contains detailed information on over 2000 published CUAs through 2008. We analyzed the extent to which the studies adhered to recommended practices for conducting and reporting cost-effectiveness analyses. We also recorded whether the studies contained information on diagnostic test accuracy and costs, and whether any account was taken of potential benefits or harms of testing that are unrelated to subsequent treatment, such as the reassurance value of testing. **Results:** We identified 141 published CUAs pertaining to diagnostic laboratory testing published through 2008 which contained 433 separate incremental cost-effectiveness ratios. Prior to 2000, there were only 20 CUAs published, but the number averaged 13.4 annually thereafter. Most studies focused on hematology/oncology (n = 42, 30%) and obstetrics/gynecology (n = 36, 26%) applications. Approximately 63% (n = 89) of studies clearly reported information about the accuracy of the test, but only 10% (n = 14) mentioned test safety or possible risks. A small number (n = 10, 7%) mentioned or considered the potential value or harm of testing unrelated to treatment consequences. Over 55% of the reported incremental cost-effectiveness ratios (ICERs) were either dominant (more quality-adjusted life years for less cost), or below $50,000 per quality-adjusted life years gained (in 2008 US dollars). **Conclusion:** The number of CUAs investigating laboratory diagnostic testing has increased substantially with applications to diverse clinical areas. The literature reveals many areas in which testing represents good value for money. The vast majority of studies have not considered preferences for test information unrelated to treatment consequences.

Keywords: cost-effectiveness, cost-utility, diagnostics, laboratory tests, QALY.

Introduction

The introduction of diagnostic laboratory techniques in clinical practice has accelerated in recent decades [1]. The new techniques have various aims: increasing certainty about the presence or absence of disease; supporting the clinical management of patients; assessing prognosis; and/or monitoring the clinical course of disease. The emergence of these tests provides new options for clinicians and patients, though it has also raised concerns about inappropriate utilization [2,3]. Payers worldwide are asking whether the benefits of new diagnostic laboratory testing are worth the costs and, if so, in what circumstances [4]. The push to demonstrate value has also meant more demanding questions about whether testing directly improves patient outcomes and whether benefits of testing outweigh the risks.

Cost-effectiveness analysis (CEA) provides a standard, well-accepted methodological technique for determining whether a health intervention provides value for money, and is used to inform reimbursement and coverage decisions in many countries [5–11]. Cost-utility analysis (CUA), a type of CEA in which the benefits of an intervention are measured in terms of quality-adjusted life years (QALYs) gained, incorporates the impact of an intervention on patient's mortality and morbidity. Because CUA allows comparisons of value across diverse interventions and conditions [12,13], they are used by many health authorities worldwide to determine value for money and are recommended by many experts in health economics and outcomes research.

In conventional CUAs, test information is typically valued only to the extent that it leads individuals and clinicians to make better medical decisions [14,15]. In practice, however, patients and health-care providers may value information from a diagnostic test even when the information does not change treatment decisions [16]. A recent study, for example, found that cancer screening was favored by 87% of adults, even in the absence of effective treatment options [17]. Of course, test information may also induce harms unrelated to treatment consequences, such as increased anxiety over test results.

CUA has been applied to a wide variety of health interventions including pharmaceuticals [18], surgical procedures [19,20], and diagnostic imaging [21,22] and is recommended by many experts in health economics and outcomes research. One previous study

* *Address correspondence to:* Peter J. Neumann, Director, Center for the Evaluation of Value and Risk in Health, Professor, Tufts University School of Medicine, Institute for Clinical Research and Health Policy Studies, 800 Washington St., Tufts Medical Center, #063, Boston, MA 02111.

E-mail: pneumann@tuftsmedicalcenter.org

1098-3015/$36.00 – see front matter

doi:10.1016/j.jval.2011.05.044

reviewed various types of published cost-effectiveness analyses involving all kinds of diagnostic strategies [23]. In contrast, our study systematically reviewed the CUA literature, focusing on diagnostic laboratory testing in different diseases. We evaluated trends in the methods used to estimate cost-effectiveness, as well as the extent to which they included key features pertinent to diagnostic laboratory technologies, such as the cost and accuracy of testing. We also explored the extent to which the QALY estimates reflected preferences for test information unrelated to treatment consequences.

Methods

We analyzed data from the Tufts Medical Center CEA Registry (www.cearegistry.org) [24], a comprehensive database of over 2000 CUAs and over 5300 cost- utility ratios published in the peer-reviewed literature up to 2008. Data from the CEA Registry have been used to review CUAs pertaining to various diseases and interventions [18,21,25,26]. The protocol for searching and including studies in the Registry is described in detail elsewhere [24]. The Registry comprises a systematic review, as each article undergoes a formal systematic screening and review process. The procedure starts with a MEDLINE search by keywords "QALYs", "quality-adjusted" and "cost-utility analysis". The search is limited to English-language publications that contain original cost-utility estimates. Reviews, editorials, and purely methodological articles are excluded, as are CEAs that measure health effects in units other than QALYs.

Each selected article is independently audited by two trained reviewers for approximately 40 elements, including key study results, information characterizing study clarity and completeness, and the methodological quality of the cost and utility weights information. The data collection form was designed based on several published guidelines and recommendations for reporting CUAs [12,27–31]. All reviewers have related background in health economics and economic evaluations in health care and received both training and detailed instructions to ensure the quality of their input data. Reviewers also assign a subjective quality score for each article (from 1 [low] to 7 [high]), which is based on factors such as whether articles present: a correct computation of the incremental cost-effectiveness ratios (ICERs); a comprehensive characterization of uncertainty; an explicit specification of assumptions; and an appropriate and explicit estimation of utility weights. All ICERs are converted to 2008 US dollars using the foreign exchange rates for currency conversion and the consumer price index (CPI) to adjust for inflation.

In this review, we selected all published CUAs that evaluated one or more diagnostic laboratory test. Tests were included if they involved samples of blood, urine, or other tissues or substances in the body [32]. In addition to the regularly audited variables [18], we collected the following additional information for each analysis: 1) clinical area of application (e.g., hematology/oncology, obstetrics/gynecology, cardiovascular); 2) type of test (e.g., genetic, virology, bacteriology, and immunology); and 3) the stage in which the test is being adopted in the marketplace (i.e., innovative testing/feasibility studies, commercially available/early adoption, or standard of care/test included in treatment guidelines by relevant physician organizations). Additional variables collected included whether the study reported the cost of testing; the test accuracy (specificity/sensitivity); the consequences of inconclusive, indeterminate, or inadequate results (e.g., whether one should conduct a different test after the result of a test is inconclusive), and the potential risks associated with testing (e.g., morbidity/mortality, adverse effects). Finally, we evaluated whether the study estimated or mentioned any consideration of the potential value or harm of testing unrelated to treatment, such as the value of reassurance, anxiety generation from false positive results, or other nonhealth attributes.

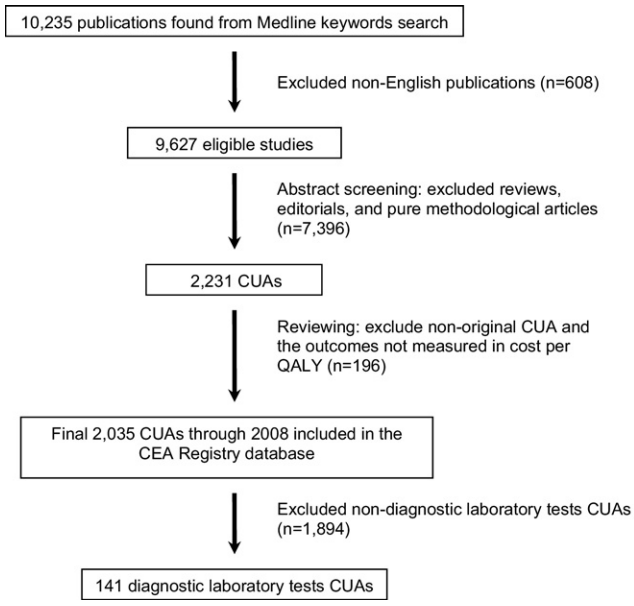

Fig. 1 – Selection of studies and review process. CEA, cost-effectiveness analysis; CUA, cost utility analyses; QALY, quality-adjusted life year.

We summarized the characteristics of all CUAs related to diagnostic laboratory tests, published through 2008. We also examined the methods used and the quality of the study. We used Student's *t*-tests to determine differences in quality scores of studies in evaluating laboratory diagnostic testing versus other types of interventions. The trend test was adopted for analyzing changes of study quality over time. Finally, we analyzed the distribution of ICERs from selected CUAs. All statistic analyses were performed using SAS 9.1 software (SAS, Cary, NC) with a definition of statistical significance at $P < 0.05$ for all tests.

Results

We identified 141 CUAs pertaining to diagnostic laboratory testing published through 2008 (Fig. 1). The first study was published in 1985; since then the number of publications has increased rapidly (Fig. 2). In the 15 years from 1985 and 1999, there were only 20 CUAs evaluating diagnostic laboratory testing; this number averaged 13.4 annually after 2000. Publication characteristics are summarized in Table 1. The analyses have been published in 86 different journals, with the most frequently publishing journals being *Annals of Internal Medicine* (n = 8 CUAs), *The American Journal of Medicine* (n = 7), and *The Journal of the American Medical Association* (n = 7).

Most of the studies (n = 88, 62.4%) were performed in the United States, followed by studies conducted in the United Kingdom (n = 13, 9.2%) and Canada (n = 6, 4.3%). The most commonly cited funding source was government agencies (n = 70, 49.6%). CUAs most frequently applied to secondary prevention defined as interventions aimed at alleviating disease by early detection, and appropriate management (n = 73, 51.8%); 27% (n = 38) were for primary prevention (interventions aimed at avoiding the onset of a symptom or disease) and 21.3% (n = 30) for tertiary prevention (interventions aimed at reducing the complications and progression of existing disease). Fifty-three studies (37.6%) investigated commercially available tests in early stages of adoption, whereas over 50% (n = 78) were for tests considered current standard of care (Table 1). The diagnos-

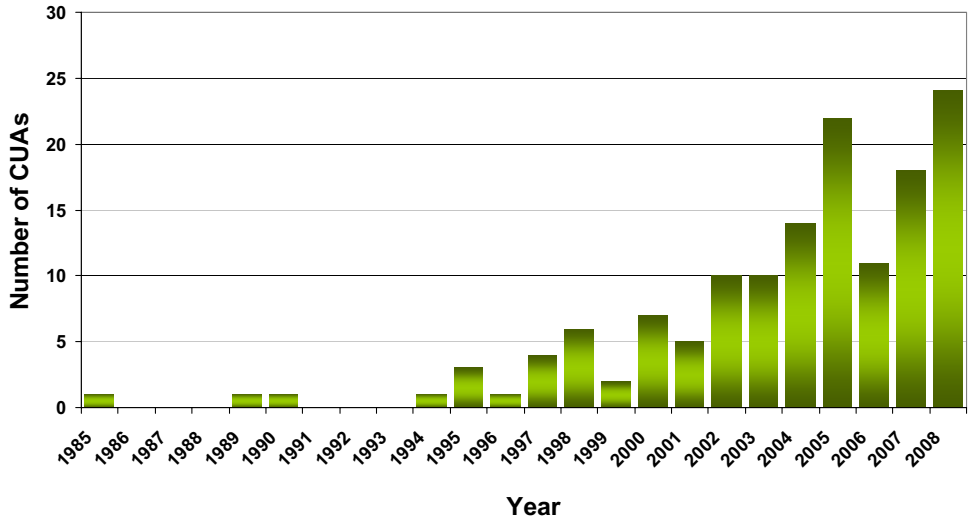

Fig. 2 – The growth in cost utility analyses (CUAs) of diagnostic laboratory tests from 1985 to 2008 (N = 141).

tic tests were applied in diverse clinical areas, including: hematology/oncology (n = 42, 29.8%), obstetrics/gynecology (n = 36, 25.5%), gastroenterology (n = 34, 24.1%), endocrinology (n = 20, 14.2%), and cardiovascular disease (n = 10, 7.1%). In terms of the types of testing, the CUAs focused most frequently on virology tests (n = 36, 25.5%), general chemistry tests (n = 30, 21.3%), and genetic testing (n = 25, 17.7%).

The vast majority of studies adhered to recommended methodological guidelines for conducting and reporting results, such as clearly presenting the intervention (99.3%); reporting the comparator and target population (99.3%); stating the study perspective (98.5%), time horizon (84.4%), currency and year of currency (85.1%); applying discounting for future costs and QALYs, (92.9%); conducting correct incremental analysis (94.3%); and performing sensitivity analyses to probe uncertainties (98.6%) (Table 2). Most studies correctly calculated ICERs; only eight CUAs (5.7%) either incorrectly reported incremental ratios or did not report relevant ratios at all.

The mean quality score of the 141 diagnostic laboratory CUAs was 4.41 (SD 1.01) compared with the quality of studies reported in other reviews using the CEA Registry data [24] was 4.27 (SD 1.06; $P = 0.1123$). No statistical significant difference was found for the study quality over time, such as whether study clearly stated target population, intervention and comparator; reported discounted rate in both costs and QALYs; and correctly calculated incremental cost-effectiveness ratios.

Approximately 63% of the CUAs reported the specificity and sensitivity of the test in question. Only 37 studies (26.2%) mentioned a subsequent strategy for dealing with inconclusive, indeterminate or inadequate results – for example, to conduct additional testing (n = 20), continue a work up if an inadequate result was obtained (n = 13), or repeat the same test (n = 8). Only 14 studies (9.9%) considered the safety or associated risks of the test either as a decrease in quality of life or as additional costs associated with complications related to testing.

A small proportion (n = 10, 7.1%) of CUAs considered and discussed the potential value or harm of testing unrelated to treatment consequences (Table 2). These studies considered various aspects related to these dimensions, including reassurance value (n = 6), unnecessary worry about false-positive results (n = 2), the benefit of reducing uncertainty (n = 1), or other attributes (n = 6). In 7 of 10 cases, these considerations of value were made in the article's discussion section, as opposed to being incorporated directly into the QALY estimate (n = 3).

Table 3 highlights studies that have considered the potential value or harms of testing unrelated to treatment consequences. These studies tended simply to mention such factors in their discussion sections, but have not attempted to quantify them. Even studies that stated that they adjusted QALY estimates for non-treatment considerations were generally vague about how estimates were constructed. For example, a CEA of expanded testing for primary HIV infection stated that "the utility of anxiety while waiting for confirmatory test results for patients with a positive screen is 0.682" [33] but did not mention how the utility weight was calculated. Another study of the cost-effectiveness of sputum cytology for lung malignancy noted that they adjusted for the effects of anxiety in the model, but did not include any details on the measurement [34].

From the 141 CUAs, we identified 433 distinct ICERs (each article may have one or more ICERs). The median ratio value per QALY gained was $36,000 (2008US$). Of the 433 ratios, 13.9% (n = 60) were dominant (QALY gains and cost-saving); whereas 43.4% (n = 188) had cost-utility estimates below the commonly used $50,000 per QALY gained threshold (Fig. 3). The proportion of ICERs that were found to be dominated (higher costs but less QALY gains) was 13.6% (n = 59). Table 4 provides specific examples selected from studies with a quality score of 5 or higher. For instance, the reported ICER of biennial fecal occult blood test in the United Kingdom was $4,700/QALY (2008US$) [35]. Compared with current hepatitis B surface antigen assays in routine donor testing in European recipients of blood transfusions, new enhanced-sensitivity hepatitis B surface antigen assays resulted in $680,000 for each additional QALY [36].

Discussion

We reviewed all published CUAs of diagnostic laboratory tests over the past 25 years to examine the study methodology and the extent to which such testing has been found to reflect a cost-effective use of resources. Our review also attempts to highlight specific areas where standardization could result in improved quality of diagnostic testing CEAs, and to monitor trends in the quality of the available published studies. We found an increasing number of published CUAs devoted to diagnostic laboratory testing in recent years, which likely reflects the rising importance of such testing in modern medicine, as well as concerns about increased utilization and overall value.

Table 1 – Characteristics of CUAs of diagnostic laboratory technology (N = 141).

	Number	Percent
Major publishing journal		
Annals of Internal Medicine	8	5.7%
American Journal of Medicine	7	5.0%
Journal of American Medical Association	7	5.0%
Transfusion	5	3.5%
Alimentary Pharmacology & Therapeutics	4	2.8%
American Journal of Obstetrics and Gynecology	4	2.8%
Pediatrics	4	2.8%
Study country		
United States	88	62.4%
United Kingdom	13	9.2%
Canada	6	4.3%
The Netherlands	5	3.5%
Other/multi-countries	29	20.6%
Funding source*		
Government	70	49.6%
Foundation	13	9.2%
Pharmaceutical/medical device companies	13	9.2%
Healthcare organization	5	3.5%
None	9	6.4%
Other/could not be determined	44	31.2%
Prevention stage		
Primary	38	27.0%
Secondary	73	51.8%
Tertiary	30	21.3%
Stage of adoption		
New test, feasibility study or test not commercially available	10	7.1%
Early adoption, commercially available but not standard of care	53	37.6%
Standard of care/test included in treatment guidelines by relevant physician organization	78	55.3%
Clinical area of application		
Hematology/oncology	42	29.8%
Obstetrics/gynecology	36	25.5%
Gastroenterology	34	24.1%
Endocrinology	20	14.2%
Cardiovascular	10	7.1%
Other	77	54.6%
Type of testing		
Virology	36	25.5%
General chemistry	30	21.3%
Genetic	25	17.7%
General immunology	21	14.9%
Immunohematology/ immunohistochemistry	18	12.8%
Other	39	27.7%

* Each article may be sponsored by multiple funding sources.

The distribution of ICERs of diagnostic laboratory testing is similar to that of other types of medical interventions [26]. More than half of reported ratios have a favorable cost-effectiveness profile using conventional benchmarks for CUAs, suggesting that many applications of testing evaluated and reported in economic evaluations may represent good value for money.

The CUAs mostly targeted hematology/oncology, OB/GYN, endocrine disorders, and cardiovascular diseases. Approximately half of the tests involved simple virology and chemistry techniques, which typically have relatively low unit costs. Nevertheless, more expensive genetic testing has also emerged as an important tool to inform treatment decisions, such as testing of BRCA1/BRCA2 for breast and ovarian cancer and testing of HLA-B*5701 to guide initial therapy for HIV. Genetic testing was evaluated in approximately 18% of the CUAs we reviewed. Given ongoing development of "personalized medicine" tied to genetic markers, this category may well expand in the future [37].

We found that the CUAs devoted to diagnostic laboratory testing performed reasonably well on many quality metrics. For ex-

Table 2 – Methods used in CUAs for diagnostic laboratory tests (n = 141).

	Number	Percent
Clear presentation of		
Intervention	140	99.3%
Comparator	140	99.3%
Target population	140	99.3%
Time horizon		
Lifetime	76	53.9%
Other	43	30.5%
Not stated	22	15.6%
Perspective		
Societal	47	33.3%
Health-care payer	92	65.2%
Not stated/could not be determined	2	1.4%
Discounting		
Cost or QALY only	9	6.4%
Both cost and QALY	116	82.3%
Not needed	6	4.3%
Not stated/could not be determined	10	7.1%
Currency and year		
Yes, reported	120	85.1%
Incremental analysis		
Correct	133	94.3%
Incorrect	5	3.6%
Could not be determined	3	2.1%
Sensitivity analysis		
Yes, reported	139	98.6%
Accuracy, safety, and cost of the test		
Considered the accuracy (specificity/sensitivity) of the test?	89	63.1%
Considered inconclusive/indeterminate/ inadequate results?	37	26.2%
Conduct additional test	20	54.1%
Continue work up	13	35.1%
Repeat the same test	8	21.6%
Considered the safety or risk of the test?	14	9.9%
Morbidity/mortality	5	35.7%
Complication or adverse effects	8	57.1%
Reported costs of the test?	121	85.8%
Value of test unrelated to treatment follow-up		
Considered the value of the test unrelated to treatment consequences?*	10	7.1%
Reassurance value	6	4.3%
Unnecessary worry about false-positive	2	1.4%
Ambiguity	1	0.7%
Other nonhealth attributes	6	4.3%
The information was located in the		
QALY estimation	3	2.1%
Discussion	7	5.0%

CUA, cost utility analysis; QALY, quality-adjusted life year.
* Not mutually exclusive.

Table 3 – How articles considered the value of testing unrelated to treatment consequences.

Article	Diagnostic laboratory test	Consideration	Location in CUA	Notes
Kievit et al. 1990[48]	Carcinoembryogenic antigen monitoring	Other nonhealth attributes	QALY estimation	"*. . .include the distress caused by false-positive test results and the fears associated with early detection of an incurable recurrence as well as the morbidity and mortality related to operations performed on account of either false- or true-positive results.*"
Raab et al. 1997[34]	Sputum cytology for lung malignancy	Reassurance, ambiguity	QALY estimation	"*Adjustments for quality of life are modeled for (1) the diagnosis of cancer, (2) test-related anxieties, and (3) test complications. . .*"; "*. . .we account for the effects of anxiety before, during and after the test and the pain and inconvenience of the test. Test result-related adjustments refer to concerns patients may have with the uncertainties of the test result (false-negative diagnosis or false-positive diagnosis) and changes in health consequent to a test complication.*"
Coco et al. 2005[33]	1) HIV-1 RNA assay; 2) p24 antigen EIA; 3) third-g	Reassurance, other nonhealth attributes	QALY estimation	"*Utility of anxiety while waiting for confirmatory test results for patients with a positive screen is 0.682.*"
Homberger et al. 2005[49]	21-Gene RT-PCR assay	Reassurance, other nonhealth attributes	Discussion	"*We deliberately omitted a number of factors that may further influence the clinical utility and economics of the assay. . .How the assay might affect patients' attitudes about their decisions has not been assess prospectively,. . .*"
Walensky et al. 2005[50]	EIA test	Reassurance, other nonhealth attributes	Discussion	"*. . .although the model included quality-of-life estimates for health states, the short-term anxiety and fear over the several days that the patient awaits test results is difficult to capture adequately when looking at total life years as the clinical endpoints. . .*"
Ball et al. 2007[51]	Pregnancy associated plasma protein A, free B-HCG, inhibin A, estriol, maternal alfafetoprotein.	Reassurance	Discussion	"*. . .It allows earlier reassurance if the results return normal. . .*"
Kobayashi et al. 2007[1]	PSA	Unnecessary worry in false-positive results, reassurance	Discussion	"*. . .It is likely that participation in the screening program leads to impairment of QOL for some men (e.g. those with false positive PSA test results). . .Satisfaction with their decision to participate*"
Killie et al. 2007[52]	Anti-HPA 1a antibodies determination	Other nonhealth attributes	Discussion	"*One may argue that women who are informed that they are at risk of having a child with severe neurological complications may develop anxiety. . .*"

(continued)

Table 3 – Continued.

Article	Diagnostic laboratory test	Consideration	Location in CUA	Notes
Nielsen M et al. 2007[53]	Genetic testing for MUTYH mutations	Other nonhealth attributes	Discussion	*". . .[W]e did not include the psychosocial aspects of genetic counseling in our analysis. People may experience changes in their functional emotional or social status after learning their genetic predisposition.". . ."because of difficulties in measuring non-health benefits. There is a need for further research on the psychosocial impact of genetic services within a health economics context"*
Smith et al. 2007[54]	DNA amplification assay for *Neisseria gonorrhoeae* and *Chlamydia trachomatis*)	Unnecessary worry in false-positive results	Discussion	*"We did not consider the possible harms of more frequent screening: adverse effects of false positive screening on patients and partners, the inconvenience of obtaining test specimens. . ."*

EIA, enzyme immunoassay; B-HCG, beta human chorionic gonadotropin; HPA, human platelet antigens; MUTYH, mutY Homolog (*Escherichia coli*); QALY, quality-adjusted life year; QOL, quality of life; PSA, prostate-specific antigen;

ample, the average quality score for the CUAs was compared with the score for other intervention types [24]. Almost all studies explicitly stated the comparator, target population, time frame, and study perspective; and most studies conducted sensitivity analyses to examine the robustness of results.

Still, we identified several areas that would benefit from increased standardization. For example, researchers, reviewers, and publishers should make sure that studies evaluating diagnostic laboratory technologies clearly state the accuracy and costs of tests, and how often sampling or processing is not adequate and what remedial steps to follow.

About one-third of CUAs did not consider or report the accuracy of the test under investigation. Most of these studies seemed to have implicitly assumed perfect accuracy of the tests. This may be one reason why those studies did not consider the value or harm unrelated to the treatment consequences, such as possible anxiety from a false-positive result. For studies that did report test accuracy, fewer than half considered a follow-up procedure for indeterminate results. As a result, the risks and costs attributable to follow-up management may have been underestimated. Invasive diagnostic technologies, such as amniocentesis, chorionic villus sampling, or genetic testing, may sometimes be associated with risks that can vary by the operator, the patient, or the test itself. However, CUAs have rarely considered such factors.

Our study also found that few CUAs have considered preferences unrelated to treatment effects. Because they ignore such

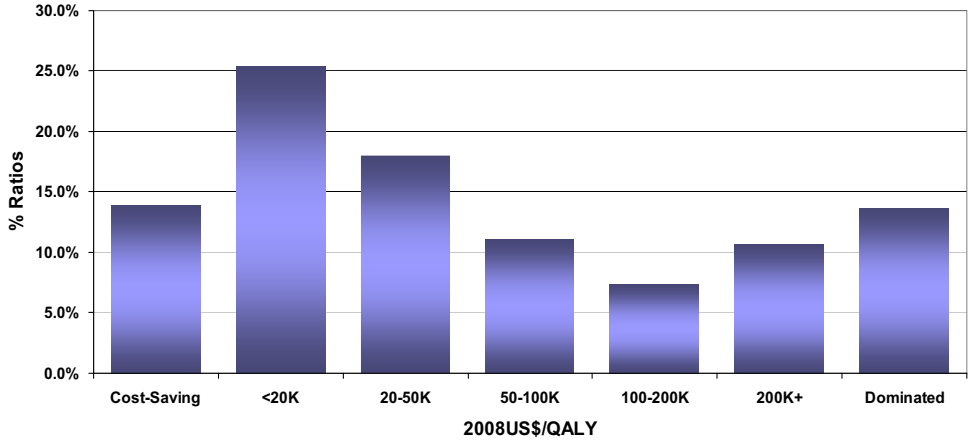

Fig. 3 – Distribution of incremental cost-effectiveness ratios for diagnostic laboratory tests (N = 433). Costs are given in 2008 US dollars (2008US$). QALY, quality-adjusted life year.

Table 4 – Selected cost-effectiveness results for diagnostic laboratory technology.

Intervention vs comparator in target population	2008US$/QALY	Reference
QFT alone vs TST followed by QFT (TST/QFT) in immunocompetent individuals aged 20 who have had contact with sputum-smear-positive pulmonary tuberculosis patients in Japan.	Cost saving	Kowada et al. 2008[55]
Screening for alpha-adducin Gly460Trp variant, and initiate diuretic therapy if positive for variant vs no screening and no diuretic treatment to any of the cohort in 65-year-old white men and women treated for hypertension, not receiving diuretics.	Cost saving	Meckley et al. 2006[56]
Biennial fecal occult blood test vs no screening in individuals aged 60–69 without polyps or cancer through to the development of adenomatous polyps and malignant carcinoma and subsequent death in the general population of England.	4,700	Tappenden et al. 2007[35]
Universal newborn screening by tandem mass spectrometry (MS/MS) for MCADD (simulated clinical course through age 20) vs no universal screening in neonates.	6,800	Venditti et al. 2003[57]
Aggressive targeted screening: human papillomavirus test using polymerase chain reaction vs no screening in HIV-infected women with CD4 of 200–500/mm^3 on highly active antiretroviral therapy.	15,000	Goldie et al. 2001[58]
Radical prostatectomy for all patients vs selection-based management policy using DNA-ploidy as an experimental marker (prostatectomy for nondiploid result; monitoring for diploid result) in male patients diagnosed with moderately differentiated (Gleason sum score 5–7) prostate cancer, age 60.	30,000	Calver et al. 2003[59]
IHC test: 1-year adjuvant trastuzumab for IHC +3 patients; standard care for all other patients vs standard care in 55-year-old female with early breast cancer completely excised and after 4 cycles of chemotherapy.	52,000	Lidgren et al. 2008[60]
Pap test and human papilloma virus test every 2 years until the age of 75 vs Pap test every 2 years until the age of 100 in hypothetical cohort of US women, age 20.	88,000	Mandelblatt et al. 2002[61]
Screening for hepatitis C within prison population, possible follow-up screening in community for those positive in prison screening vs usual care – no prison screening, only symptomatic testing within community in prison population of England and Wales.	110,000	Sutton et al. 2008[62]
New enhanced-sensitivity HBsAg assays in routine donor testing vs current HBsAg assays in routine donor testing in European recipients of blood transfusions.	680,000	Pereira 2003[36]
Routine voluntary HIV counseling, testing, and referral screening vs current practice in US general population: prevalence of undiagnosed HIV 0.1%, annual incidence of 0.01%.	Dominated	Paltiel et al. 2005[63]
Optical immunoassay alone with antibiotic treatment for positive results vs throat culture with antibiotic treatment for positive results in adult patients with suspected group A beta-hemolytic *Streptococcus* pharyngitis.	Dominated	Neuner et al. 2003[64]

Costs are given in 2008 US dollars (2008US$).
Source: Tufts Medical Center, the Cost-Effectiveness Analysis Registry (www.cearegistry.org).[24]
HBsAg, hepatitis B surface antigen; IHC, immunohistochemical; MCADD, medium chain acyl-CoA dehydrogenase deficiency; QFT, QuantiFERON-TB Gold; TST, tuberculin skin test.

factors, existing CEAs of diagnostic testing may underestimate the true value of testing. Clinicians may prefer tests in part because they reduce diagnosis uncertainty, even if patients do not actually fare better after been tested [38]. Patients may favor testing due to various factors, including their preference for reduced uncertainty about their disease state or prognosis [14,39–42]. One recent survey found that most of respondents would be inclined to take tests predictive of future cancer, Alzheimer's disease, or arthritis, even in the absence of immediate treatment consequences [43]. Study respondents valued testing for a host of nonmedical reasons (e.g., they would live their lives differently or get their finances in order) [43].

Future research should focus more on how to value diagnostic testing beyond their influences in therapy. One recent study found that awaiting breast biopsy provided greater stress for woman in terms of anxiety and perceived stress than awaiting much riskier invasive treatment of cancer [44]. To avoid over- or underestimating the value of diagnostic laboratory tests, we recommend that the value of test information unrelated to treatment consequences should be considered in future cost-effectiveness evaluation. How to value the information or value "knowing for the sake of knowing" will be an area of special interest and challenges for the future research [39,43]. At the very least, authors of CEAs of diagnostic laboratory tests should qualitatively discuss the potential benefits and harms, such as the "value of knowing" or increased anxiety associated with testing, that were not captured in the analysis. Beyond that, investigators might include the implications of test information in the health state descriptions used in their utility assessments; however, how to accomplish this without unduly influencing respondent preferences will be a challenge. Additionally, more research from contingent valuation studies on the willingness to pay for test information even where there are no direct treatment implications would be helpful [43].

Health technology assessment (HTA) agencies worldwide are intensifying their scrutiny of diagnostic laboratory tests and are increasingly using information from CEAs. Our study suggests that HTA organizations might better account for individuals' "value of knowing" from such tests. A key policy question will be whether and to what extent patients should pay themselves for such nonhealth outcomes.

Our study has several limitations. First, studies were included in the registry only if they met several peer-reviewed, English-language, original CUAs presenting a cost per QALY ratio available in Medline. Notwithstanding the limitations of our search strategy and inclusion criteria, the sample of studies included in our review is fairly large and diverse and likely is a broad representative of studies available in this field. Previous evidence suggests that manual searches and searches of databases other than Medline have a limited incremental yield [45]. A study of economic evaluations of laboratory diagnostics beyond CUAs would be helpful in the future but also presents challenges in terms of comparing studies with disparate endpoints.

Second, our measure of study quality was based on a 7-point subjective Likert scale agreed upon by two expert readers and, therefore, might have been subject to reader bias or not sensitive enough to capture some subtle quality differences. However, the

scale has been analyzed and reported on in several previous publications, which found that the quality score was correlated with factors such as the methodology used to conduct and report results of sensitivity analyses, as well as the quality and experience of journals reporting results from economic evaluations [18,21,22,26,46]. We did not judge the appropriateness of the comparator used in the incremental cost-effectiveness ratios. Furthermore, our results could be influenced by a publication bias (in which positive results are published more often) that is found in the general CUAs literature [47].

Finally, by summarizing and analyzing all articles published over a 20-year period as a group, slight changes and trends in publications, methodology, and/or inclusion of key parameters could have been missed. The relatively low number of publications per year and variability in the quality of the studies, however, makes it impossible to find statistical significance regarding finer details.

In conclusion, CEA of laboratory diagnostic testing has increased rapidly in recent years and the information can help decision makers in their coverage and reimbursement decisions. Many of the CUAs have adhered to key recommended protocols for the field. Very few studies, however, have considered the potential value or harm of testing unrelated to treatment consequences.

Source of financial support: This study was funded by a foundation grant from the Institute for Health Technology Studies (InHealth).

REFERENCES

[1] Kobayashi T, Goto R, Ito K, Mitsumori K. Prostate cancer screening strategies with re-screening interval determined by individual baseline prostate-specific antigen values are cost-effective. Eur J Surg Oncol 2007;33:783–9.

[2] Thomson Reuters. Waste in the US Healthcare System Pegged at $700 Billion in Report from Thomson Reuters. Available from: http://thomsonreuters.com/content/press_room/tsh/waste_US_healthcare_system [Accessed July 8, 2011].

[3] van WC, Naylor CD. Do we know what inappropriate laboratory utilization is? A systematic review of laboratory clinical audits. JAMA 1998;280:550–8.

[4] Gluud C, Gluud LL. Evidence based diagnostics. BMJ 2005;330:724–6.

[5] Neumann PJ. Using Cost-Effectiveness Analysis to Improve Health Care. New York, New York: Oxford University Press; 2005.

[6] Neumann PJ, Rosen AB, Weinstein MC. Medicare and cost-effectiveness analysis. N Engl J Med 2005;353:1516–22.

[7] Garber AM. Cost-effectiveness and evidence evaluation as criteria for coverage policy. Health Aff (Millwood) 2004; Suppl Web Exclusives: W4-284-96.:W4-96.

[8] Harris A, Buxton M, O'Brien B, Drummond FR. Using economic evidence in reimbursement decisions for health technologies: experience of 4 countries. Expert Rev Pharmacoecon Outcomes Res 2001;1:7–12.

[9] Ramsey SD, Veenstra DL, Garrison LP, Jr.et al. Toward evidence-based assessment for coverage and reimbursement of laboratory-based diagnostic and genetic tests. Am J Manag Care 2006;12:197–202.

[10] O'Malley SP. The Australian experiment: the use of evidence based medicine for the reimbursement of surgical and diagnostic procedures (1998-2004). Aust New Zealand Health Policy 2006;3:3.

[11] Drummond MF, Mason AR. European perspective on the costs and cost-effectiveness of cancer therapies. J Clin Oncol 2007;25:191–5.

[12] Drummond MF, O'Brien B, Stoddart GL, Torrance GW. Methods for the Econonmic Evaluation of Health Care Programmes. Oxford, England: Oxford University Press; 2003.

[13] Weintraub WS, Cohen DJ. The limits of cost-effectiveness analysis. Circ Cardiovasc Qual Outcomes 2009;2:55–8.

[14] Baker LC, Atlas SW, Afendulis CC. Expanded use of imaging technology and the challenge of measuring value. Health Aff (Millwood) 2008;27:1467–78.

[15] Grosse SD, Wordsworth S, Payne K. Economic methods for valuing the outcomes of genetic testing: beyond cost-effectiveness analysis. Genet Med 2008;10:648–54.

[16] Lee DW, Neumann PJ, Rizzo JA. Understanding the medical and nonmedical value of diagnostic testing. Value Health 2010;13:310–4.

[17] Schwartz LM, Woloshin S, Fowler FJ, Jr., Welch HG. Enthusiasm for cancer screening in the United States. JAMA 2004;291:71–8.

[18] Neumann PJ, Fang CH, Cohen JT. 30 years of pharmaceutical cost-utility analyses: growth, diversity and methodological improvement. Pharmacoeconomics 2009;27:861–72.

[19] Brauer CA, Neumann PJ, Rosen AB. Trends in cost effectiveness analyses in orthopaedic surgery. Clin Orthop Relat Res 2007;457:42–8.

[20] Brauer CA, Rosen AB, Olchanski NV, Neumann PJ. Cost-utility analyses in orthopaedic surgery. J Bone Joint Surg Am 2005;87:1253–9.

[21] Otero HJ, Rybicki FJ, Greenberg D, Neumann PJ. Twenty years of cost-effectiveness analysis in medical imaging: are we improving? Radiology 2008;249:917–25.

[22] Otero HJ, Rybicki FJ, Greenberg D, Mitsouras D, Mendoza JA, Neumann PJ. Cost-effective diagnostic cardiovascular imaging: when does it provide good value for the money? Int J Cardiovasc Imaging 2010;26:605–12.

[23] Redekop WK. Cost-effectiveness analyses of diagnostic strategies: a literature survey using the NHS Economic Evaluation Database. Expert Rev Pharmacoecon Outcomes Res 2006;6:41–8.

[24] Tufts Medical Center. Cost-Effectiveness Analysis Registry. Available from: http://www.cearegistry.org. 2010. [Accessed June 15, 2011].

[25] Talmor D, Greenberg D, Howell MD, Lisbon A, Novack V, Shapiro N. The costs and cost-effectiveness of an integrated sepsis treatment protocol. Crit Care Med 2008;36:1168–74.

[26] Greenberg D, Earle C, Fang CH, Eldar-Lissai A, Neumann PJ. When is cancer care cost-effective? A systematic overview of cost-utility analyses in oncology. J Natl Cancer Inst 2010;102:82–8.

[27] Economic analysis of health care technology. A report on principles. Task Force on Principles for Economic Analysis of Health Care Technology. Ann Intern Med 1995;123:61–70.

[28] Drummond MF, Richardson WS, O'Brien BJ, Levine M, Heyland D. Users' guides to the medical literature. XIII. How to use an article on economic analysis of clinical practice. A. Are the results of the study valid? Evidence-Based Medicine Working Group. JAMA 1997;277:1552–7.

[29] Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ 1996;313:275–83.

[30] Menon D, Schubert F, Torrance GW. Canada's new guidelines for the economic evaluation of pharmaceuticals. Med Care 1996;34(12 Suppl):DS77–DS86.

[31] O'Brien BJ, Heyland D, Richardson WS, Levine M, Drummond MF. Users' guides to the medical literature. XIII. How to use an article on economic analysis of clinical practice. B. What are the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. JAMA 1997;277:1802–6.

[32] U.S. National Library of Medicine. Laboratory Tests. MedlinePlus. 2010.Ref Type: Electronic Citation

[33] Coco A. The cost-effectiveness of expanded testing for primary HIV infection. Ann Fam Med 2005;3:391–9.

[34] Raab SS, Hornberger J. The effect of a patient's risk-taking attitude on the cost effectiveness of testing strategies in the evaluation of pulmonary lesions. Chest 1997;111:1583–90.

[35] Tappenden P, Chilcott J, Eggington S, Patnick J, Sakai H, Karnon J. Option appraisal of population-based colorectal cancer screening programmes in England. Gut 2007;56:677–84.

[36] Pereira A. Health and economic impact of posttransfusion hepatitis B and cost-effectiveness analysis of expanded HBV testing protocols of blood donors: a study focused on the European Union. Transfusion 2003;43:192–201.

[37] Ginsburg GS, McCarthy JJ. Personalized medicine: revolutionizing drug discovery and patient care. Trends Biotechnol 2001;19:491–6.

[38] DeKay ML, Asch DA. Is the defensive use of diagnostic tests good for patients, or bad? Med Decis Making 1998;18:19–28.

[39] Asch DA, Patton JP, Hershey JC. Knowing for the sake of knowing: the value of prognostic information. Med Decis Making 1990;10:47–57.

[40] Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making 1991;11:88–94.

[41] Shogren J. Economics of diet and health: research challenges. Section C, 2, 117-127. 2005. Food Economics: Acta Agriculture Scandinavica. Ref Type: Report

[42] Viscusi W, Magat W, Huber J. Communication of ambiguous risk information. Theory and Decision 1991;31:159–73.

[43] Neumann PJ, Cohen J.T., Hammitt J.K. et al. Willingness to pay for predictive tests with no immediate treatment implications: a survey of U.S. residents. Health Economics 2010 Dec 28 [Epub ahead of print].

[44] Flory N, Lang EV. Distress in the radiology waiting room. Radiology 2011;260:166–73.

[45] Sassi F, Archard L, McDaid D. Searching literature databases for health care economic evaluations: how systematic can we afford to be? Med Care 2002;40:387–94.

[46] Meckley LM, Greenberg D, Cohen JT, Neumann PJ. The adoption of cost-effectiveness acceptability curves in cost-utility analyses. Med Decis Making 2010;30:314–9.

[47] Bell CM, Urbach DR, Ray JG et al. Bias in published cost effectiveness studies: systematic review. BMJ 2006;332(7543):699–703.

[48] Kievit J, van d, V. Utility and cost of carcinoembryonic antigen monitoring in colon cancer follow-up evaluation. A Markov analysis. Cancer 1990;65:2580–7.
[49] Hornberger J, Cosler LE, Lyman GH. Economic analysis of targeting chemotherapy using a 21-gene RT-PCR assay in lymph-node-negative, estrogen-receptor-positive, early-stage breast cancer. Am J Manag Care 2005;11:313–24.
[50] Walensky RP, Weinstein MC, Kimmel AD et al. Routine human immunodeficiency virus testing: an economic evaluation of current guidelines. Am J Med 2005;118:292–300.
[51] Ball RH, Caughey AB, Malone FD et al. First- and second-trimester evaluation of risk for Down syndrome. Obstet Gynecol 2007;110:10–7.
[52] Killie MK, Kjeldsen-Kragh J, Husebekk A, Skogen B, Olsen JA, Kristiansen IS. Cost-effectiveness of antenatal screening for neonatal alloimmune thrombocytopenia. BJOG 2007;114:588–95.
[53] Nielsen M, Hes FJ, Vasen HF, van den Hout WB. Cost-utility analysis of genetic screening in families of patients with germline MUTYH mutations. BMC Med Genet 2007;8:42.
[54] Smith KJ, Cook RL, Roberts MS. Time from sexually transmitted infection acquisition to pelvic inflammatory disease development: influence on the cost-effectiveness of different screening intervals. Value Health 2007;10:358–66.
[55] Kowada A, Takahashi O, Shimbo T, Ohde S, Tokuda Y, Fukui T. Cost effectiveness of interferon-gamma release assay for tuberculosis contact screening in Japan. Mol Diagn Ther 2008;12:235–51.
[56] Meckley LM, Veenstra DL. Screening for the alpha-adducin Gly460Trp variant in hypertensive patients: a cost-effectiveness analysis. Pharmacogenet Genomics 2006;16:139–47.
[57] Venditti LN, Venditti CP, Berry GT et al. Newborn screening by tandem mass spectrometry for medium-chain Acyl-CoA dehydrogenase deficiency: a cost-effectiveness analysis. Pediatrics 2003;112:1005–15.
[58] Goldie SJ, Freedberg KA, Weinstein MC, Wright TC, Kuntz KM. Cost effectiveness of human papillomavirus testing to augment cervical cancer screening in women infected with the human immunodeficiency virus. Am J Med 2001;111:140–9.
[59] Calvert NW, Morgan AB, Catto JW et al. Effectiveness and cost-effectiveness of prognostic markers in prostate cancer. Br J Cancer 2003;88:31–5.
[60] Lidgren M, Jonsson B, Rehnberg C, Willking N, Bergh J. Cost-effectiveness of HER2 testing and 1-year adjuvant trastuzumab therapy for early breast cancer. Ann Oncol 2008;19:487–95.
[61] Mandelblatt JS, Lawrence WF, Womack SM et al. Benefits and costs of using HPV testing to screen for cervical cancer. JAMA 2002;287: 2372–81.
[62] Sutton AJ, Edmunds WJ, Sweeting MJ, Gill ON. The cost-effectiveness of screening and treatment for hepatitis C in prisons in England and Wales: a cost-utility analysis. J Viral Hepat 2008;15:797–808.
[63] Paltiel AD, Weinstein MC, Kimmel AD et al. Expanded screening for HIV in the United States--an analysis of cost-effectiveness. N Engl J Med 2005;352:586–95.
[64] Neuner JM, Hamel MB, Phillips RS, Bona K, Aronson MD. Diagnosis and management of adults with pharyngitis. A cost-effectiveness analysis. Ann Intern Med 2003;139:113–22.